4Б

УДК 004.942

А.В. Бухановский, С.В. Иванов, Ю.И. Нечаев
Санкт-Петербургский государственный университет
информационных технологий, механики и оптики
197101, Санкт-Петербург, пр. Кронверкский, 49, avb_mail@mail.ru

Особенности планирования эксперимента при моделировании экстремальных ситуаций в интеллектуальной системе исследовательского проектирования

A.V. Boukhanovsky, S.V. Ivanov, Yu.I. Nechaev
St. Petersburg State University of Information Technologies, Mechanics and Optics
197101, Russia, Saint Petersburg, Kronverkskiy pr., 49

Features of Planning Experiment at Modeling of Extreme Situations in Intelligence System of Research Designing

А.В. Бухановський, С.В. Іванов, Ю.І. Нєчаєв
Санкт-Петербурзький державний університет
інформаційних технологій, механіки і оптики
197101, Росія, пр. Кронверкський, 49, м. Санкт-Петербург, avb_mail@mail.ru

Особливості планування експерименту при моделюванні екстремальних ситуацій в інтелектуальній системі дослідного проектування

Обсуждаются вопросы планирования эксперимента в интеллектуальных системах исследовательского проектирования морских судов. Особое внимание обращается на решение проблемы планирования эксперимента при реализации принципа конкуренции в условиях неопределенности и неполноты исходной информации.
Ключевые слова: интеллектуальная система, планирование эксперимента, принцип конкуренции

Questions of planning experiment in intelligence systems of research designing of marine ships are discussed. The special attention addresses on the decision of a problem of planning experiment at realization of a competition principle in the conditions of uncertainty and incompleteness of the initial information.
Key words: intellectual system, design of experiments, the principle of competition.

Обговорюються питання планування експерименту в інтелектуальних системах дослідного проектування морських суден. Особлива увага звертається на розв'язання проблеми планування експерименту при реалізації принципу конкуренції в умовах невизначеності і неповноти вихідної інформації.
Ключові слова: інтелектуальна система, планування експерименту, принцип конкуренції.

1 Концепция моделирования экстремальных ситуаций в задачах исследовательского проектирования

При реализации интеллектуальных технологий в интеллектуальных системах (ИС) исследовательского проектирования на базе высокопроизводительных средств вычис-

лений учитываются общие принципы и структура информационной модели, обеспечивающей анализ и прогноз исследуемых ситуаций. Такая модель разрабатывается в рамках нечеткого логического базиса и формализованных методов оценки динамики взаимодействия судна в различных условиях эксплуатации, в том числе и в экстремальных ситуациях [1-15].

В настоящей статье рассматривается приложение разработанной концепции интерпретации текущих ситуаций в сложных динамических средах [9]. Объектом исследования выбрана область практических приложений, связанная с контролем динамики судна на волнении на этапе исследовательского проектирования (рис. 1).

Задача выбора альтернатив и принятия решений в нечеткой среде связана с построением варианта (выбора альтернатив), удовлетворяющего заданным ограничениям и являющегося способом достижения цели. В процессе решения задачи, представленной кортежем [7], [11]:

$$< A, W, S, U >, \tag{1}$$

где A – множество альтернатив; W – характеристики внешней среды; S – система предпочтений; U – действие, необходимо выполнить определенную операцию над множеством альтернатив, установив наиболее предпочтительную из них.

Решение задачи (1) ведется с использованием нечетких графов и нечетких отношений [2]. Концептуальный базис нечеткой среды моделирования и планирования эксперимента основан на использовании парадигмы обработки информации в мультипроцессорной вычислительной среде [8] и достижений в области интеллектуальных технологий XXI века [7], [11]. Принципы преобразования информации в трудноформализуемых средах сформулированы в работах [2], [9].

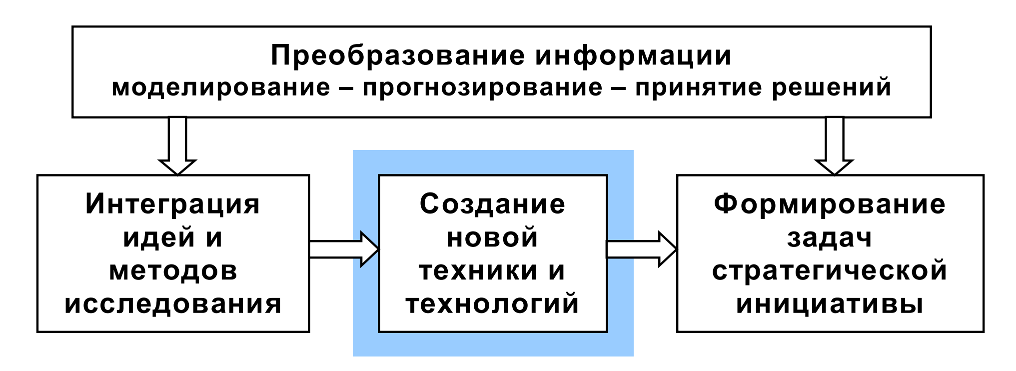


Рисунок 1 – Информационная модель создания новой техники и технологий в ИС исследовательского проектирования

Формализованное ядро системы интеллектуальной поддержки процессов построения и использования моделей знаний при анализе текущих ситуаций реализуется в рамках нечеткого логического базиса. Фундаментальной основой такой интерпретации является концепция нечетких целей и ограничений [2], [7], [11].

Повышение эффективности функционирования процедурной компоненты базы знаний интеллектуальной системы (ИС) достигается за счет использования принципа конкуренции и формализации процесса обработки нечеткой информации [1] в высокопроизводительных вычислительных средах. Другими принципами повышения эффективности ИС новых поколений являются принцип открытости, принцип сложности и принцип нелинейной самоорганизации [9]. Реализация указанных принципов осуществляется в рамках концепции мягких вычислений [15], интегрирующей нечеткую логику [1], ИНС [7] и генетический алгоритм [4].

2 Реализации принципа конкуренции при анализе альтернативных решений в ИС исследовательского проектирования

Интерпретация текущих ситуаций в сложной динамической среде осуществляется в рамках принципа конкуренции [8], обеспечивающего выбор предпочтительной вычислительной технологии обработки информации методами теории катастроф [9] на основе концепции мягких вычислений [15].

Рассмотрим прикладную теорию: концепцию принципа конкуренции (Competition Principle Concept – CPC) как систему вида [8], [9]:

$$CPC \Rightarrow S: S = \langle R, P, Q \rangle, \tag{2}$$

где S – система, отображающая теорию CPC; R – классы всех операций и математических зависимостей, описываемых CPC; P – совокупность операций на множествах R; Q – множества отношений между элементами класса R.

На основе отображения (2) представим *компьютерную интерпретацию* теории CPC как

$$J(CPC)\ S^*: S^* = \langle R^*,P^*,Q^* \rangle, \tag{3}$$

Здесь символ * указывает на различия между указанными множествами, определяемые особенностями компьютерной интерпретации. Условия сохранения целостности информации о динамике взаимодействия обеспечиваются путем реализации парадигмы обработки информации в мультипроцессорной вычислительной среде [8]. Необходимая производительность достигается с помощью процессоров, которые являются узлами вычислительной сети единой распределенной системы. При реализации принципа конкуренции используется бортовой многопроцессорный комплекс стандартной архитектуры [2], [7], [11].

Выбор предпочтительного решения в рамках принципа конкуренции осуществляется подсистемой поддержки принятия решений. Постановка задачи (1) реализуется на основе подхода Беллмана – Заде и метода иерархий (парных сравнений) Саати [2].

Подход Беллмана – Заде. Рассмотрим отображение $\varphi:X\to Y$, значениями которого являются элементы множеств X и Y. Это отображение понимается как реакция системы на входные воздействия $x\in X$ или как некоторые оценки выбора соответствующих альтернатив. Нечеткая цель задается в виде нечеткого подмножества универсального множества реакций Y, т.е. в виде функции принадлежности (ФП) $\mu_G:Y\to[0,1]$. Решение задачи состоит в достижении цели при заданных нечетких ограничениях. Пусть некоторая альтернатива x обеспечивает достижение цели со степенью $\mu_G(x)$ и удовлетворяет ограничениям со степенью $\mu_C(x)$. Тогда степень принадлежности альтернативы x решению задачи равна минимальному из этих чисел, т.е. ФП решения:

$$\mu_G(x) = \min\{\mu_G(x), \mu_C(x)\}. \tag{4}$$

При наличии нескольких целей и ограничений, а также с учетом их важности λ_i и ν_j нечеткое решение описывается соответствующими ФП, а неопределенность выбора альтернативы с максимальной степенью принадлежности нечеткому решению D определяется условием:

$$\max_{x\in X} \mu_D(x) = \max_{x\in X} \min\{\mu_G(x), \mu_C(x)\}. \tag{5}$$

Метод анализа иерархий. Реализация этого метода связана с выделением приоритетов (весов) признаков в целях выбора наилучших из них с использованием алгебраической теории матриц и экспертных процедур:

$$W\overline{\pi} = \lambda_{\max}\overline{\pi};$$
$$\Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1m} \\ \vdots & \pi_{ij} & \vdots \\ \pi_{n1} & \cdots & \pi_{nm} \end{bmatrix} \times \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix}, \quad (6)$$

где W – обратно симметричная матрица значений парных сравнений признаков относительно данного атрибута; $\overline{\pi}$ – нормированный вектор весов признаков; $\lambda_{\max}$ – наибольшее собственное значение матрицы W; П – результат определения глобальных приоритетов признаков $\Pi_1,\ldots,\Pi_N$; N – число признаков; π_{ij} (i=1,…,n, j=1,…,m) – относительный вес i-го признака по j-му атрибуту; g_j – относительный вес j-го атрибута.

На множестве альтернатив каждому критерию ставятся в соответствие лингвистические переменные (ЛП) с терм-множеством $T_i=\{T_{i1},\ldots,T_{it}\}$, которые характеризуют предпочтения конструктора. Базовым множеством ЛП является допустимое множество значений критерия. Оценка эффективности альтернативы по отдельному критерию осуществляется с помощью дополнительной ЛП с терм-множеством $S=\{S_1,\ldots,S_l\}$, мощность которого совпадает с T_i. Введенная таким образом ЛП инвариантна по отношению к каждому из критериев CR_i, а ее базовое множество принимает значения на отрезке [0,1]. Между двумя ЛП существует причинно-следственная связь в виде нечеткого импликативного отношения:

$$\mu_{Rij}(x_i,y):\ X_i \times Y \to [0,1],\ x_i \in X_i, y \in Y, \quad (7)$$

где X_i и Y – базовые множества значений i-го критерия и оценки альтернатив.

При функционировании ИС вычислительный эксперимент выполняется в виде следующих этапов: формирование альтернатив (сценариев экстремальных ситуаций); оптимизация и тестирование вариантов; анализ альтернатив и выбор предпочтительного решения.

На основании подхода Беллмана – Заде и метода анализа иерархий разработан алгоритм и программное средство моделирования и интерпретации экстремальных ситуаций при реализации парадигмы обработки информации в мультипроцессорной вычислительной среде в ИС исследовательского проектирования. Наряду с подходом Беллмана – Заде и методом анализа иерархий Т. Саати при интерпретации решений в системе интеллектуальной поддержки конструктора могут использоваться и другие модели. Среди них следует выделить функцию выбора, модель выбора и метод эталонов [11].

3 Особенности планирования эксперимента ИС исследовательского проектирования

При разработке сценариев взаимодействия судна с внешней средой широко используются методы планирования эксперимента. Содержательный анализ проблем моделирования, определяемых соотношением между реальным экспериментом, модельным экспериментом и теорией приводит к поиску стратегий моделирования с учетом особенностей исходной информации и производительности вычислительных средств. Формальные процедуры, лежащие в основе функционирования ИС исследовательского проектирования, предполагают широкое использование измерительной информации и методов математического моделирования. Решение прикладных задач анализа и

прогноза ведется с привлечением теории *планирования эксперимента*. Информация, получаемая в процессе нормальной эксплуатации ДО, представляет собой данные пассивного эксперимента. Активный эксперимент используется только на этапе тестирования и адаптации ИС. При использовании данных активного эксперимента проблема выбора наилучшего математического описания может быть решена классическими методами планирования. В случае пассивного эксперимента эта проблема требует специальных подходов, учитывающих адекватность описания, дисперсию оценок и вычислительные трудности из-за вырождения информационной матрицы [5]. Анализ подходов к планированию эксперимента и поиску эффективных процедур обработки информации в сложных ИС реального времени содержится в работах [2], [7], [9], [11].

Возникающие при практическом использовании ИС исследовательского проектирования задачи планирования эксперимента часто выходят за рамки классического подхода и требуют специального обобщения с учетом особенностей рассматриваемых проблем.

Рассмотрим определения, связанные с построением планов эксперимента при реализации моделей интерпретации динамических ситуаций [3], [5], [9].

Определение 1. Под планом эксперимента ξ будем понимать вероятностную меру на области планирования U с конечным носителем $\sup \xi = \{h \in U, \xi(h) \neq 0\}$. Обозначим через H_1 конечномерное подпространство некоторого пространства H ($H_1 \subseteq H$), которому принадлежит функция регрессии, и будем считать, что по условиям эксперимента могут быть использованы только планы из допустимого множества

$$J_N := \{\xi \in J_N : \mathrm{card}\,(\sup \xi) \leq n\}, \tag{8}$$

где n – число, характеризующее ограниченность ресурсов.

Определение 2. Выбор пространства оценивания $H_1 = H_\xi$ и оператора оценивания $S = S_\xi$ обеспечивает при фиксированном плане $\xi \in J_N$ нахождение наилучшего в метрике пространства H приближения $E\hat{\eta} \in H_\xi$ к произвольному элементу η из H. При этом пространство оценивания H_ξ должно обеспечивать построение оценки $\hat{\eta}$ неизвестного элемента $\eta \in H$.

Определение 3. Задача выбора процедуры $\pi = (\xi, H_1, S)$ восстановления η из H на множестве допустимых процедур π является двухкритериальной. В этих условиях поиск оптимальной процедуры π^* удобно вести на основе решения задачи оптимизации с приоритетом, учитывающим систематическую ошибку.

$$\pi^* = \mathrm{Arg}\inf B_\gamma(\pi), \tag{9}$$

где γ – вероятностная мера для осреднения систематической ошибки $B(\xi, H_1, S)$

$$B_\gamma(\xi, H_1, S) := \int_H B(\xi, H_1, S)\, d\gamma(\eta);$$

$$B(\xi, H_1, S) := dist^2(\eta, E\hat{\eta}) := \|\eta - E\hat{\eta}\|_H^2;$$

S – оператор, с помощью которого производится оценка $\hat{\eta} = \hat{\eta}(y)$ неизвестного элемента $\eta \in H$.

Определение 4. Для характеристики случайной ошибки используется функционал $\Phi(\xi, H_1, S)$ от корреляционного оператора оценки $\hat{\eta}$ (определитель $\det D(\hat{\eta})$, след $\mathrm{tr}\, D(\hat{\eta})$ или другие критерии), и тогда

$$\pi^* = Arg \inf_{\pi \in \tilde{\Pi}} \Phi(\pi), \tag{10}$$

где $\tilde{\Pi} \in \Pi$ – множество процедур π^*, являющихся решением задачи (9).

Определение 5. Задача (8) на множестве процедур

$$\xi^* \in J_N^* := \left\{ \tilde{\xi} \in J_N : \tilde{\xi} = Arg \inf B_\gamma(\pi_\xi) \right\} \tag{11}$$

рассматривается в предположении, что носитель плана sup ξ определен однозначно из решения задачи (10), но имеется свобода в выборе весов наблюдений P_j (j = 1, m). Это позволяет минимизировать случайную ошибку при следующих условиях

$$\xi^*(P) := (h_1^*, ..., h_N^*, P_1, ..., P_m);\ P^* = Arg \inf \Phi[D_{\xi^*(P)}], \tag{12}$$

где $P := \{P \in R^m;\ P_j > 0,\ j=1,m,\ \Sigma_{j=1}^{m} \}$; $D_{\xi^*(P)}$ – корреляционный оператор, определяемый как $D_\xi = M_x^{-1}$ (M_ξ – информационный оператор).

В прикладных задачах планирования эксперимента важное значение приобретает интерпретация таких свойств, как ортогональность и ротатабельность. При равноточных измерениях эти свойства приводят к независимости коэффициентов регрессии (их одинаковой точности), что имеет принципиальное значение в поисковых процедурах Бокса-Уилсона при *экстремальном планировании*.

Выбор оптимальных условий эксперимента. Задача выбора оптимальных условий эксперимента в ИС определяет надежную оценку характеристик ДО и параметров внешней среды и связана с построением нормированного дискретного плана для ДО, развивающегося во времени и пространстве.

Определение 6. Стратегия выбора оптимальных условий эксперимента характеризуется тем, что объект функционирует в режиме нормальной эксплуатации. ИС на основе анализа ситуации выбирает моменты времени и координаты точек, в которых следует производить измерения. Выполнив серию опытов при некоторых фиксированных значениях исследуемого фактора в различные моменты времени и имея модель системы, можно подсчитать нормированную информационную матрицу

$$M(\varepsilon) = \sum_{\nu=1}^{n} P_\nu M(\ell_\nu) \tag{13}$$

для дискретного плана эксперимента

$$\xi = \begin{Bmatrix} \ell_1, \dots, \ell_n \\ P_1, \dots, P_n \end{Bmatrix}, \quad \sum_{\nu=1}^{n} P_\nu = 1 \left(P_\nu \geq 0\right), \tag{14}$$

где P_ν – веса наблюдений.

Определение 7. Синтез D-оптимальных планов измерений осуществляется на основе итерационной процедуры с использованием функции

$$\delta(\ell,\xi) = \sup[M^{-1}(\xi)M(\ell)]. \tag{15}$$

Определение 8. Особым случаем идентификации динамических систем является планирование измерений в частотной области. Управление процессом измерений на основании методов планирования эксперимента позволяет свести задачу к поиску дискретного оптимального плана

$$\xi = \begin{Bmatrix} \omega_1, \dots \omega_n \\ P_1, \dots, P_n \end{Bmatrix} \tag{16}$$

с нормированной информационной матрицей

$$M(\xi) = \frac{1}{2\pi}\sum_{\nu=1}^{n}\Phi^{*}(i\omega_{\nu})\Phi^{T}(i\omega_{\nu})P_{\nu} + \frac{1}{2\pi}\sum_{\nu=1}^{n}\Phi(i\omega_{\nu})\Phi^{*T}(i\omega_{\nu})P_{\nu}, \quad (17)$$

где ω_{ν} – частота спектрального разложения стационарной случайной функции (входного процесса); P_{ν} – ординаты непрерывного спектра этой функции; звездочкой (*) помечены комплексно сопряженные частотные характеристики.

Интервал, на котором определена вероятностная мера, порожденная нормированной спектральной плотностью, характеризуется выражением

$$\Omega_{\omega} = [-\omega_{k}, +\omega_{k}]. \quad (18)$$

где $[-\omega_k, +\omega_k]$ – диапазон частот, определяющий полосу пропускания.

Спектр плана в этом случае представляет собой совокупность значений частот, а дисперсии гармоник являются весами.

Практическое применение методов планирования измерительного эксперимента реализовано в виде встроенной процедуры в системе автоматизации эксперимента на стадии исследовательского проектирования. Процедура реализует выбор оптимальных условий измерений, обеспечивающих надежную оценку характеристик ДО и параметров внешней среды, а также формализацию процедурной компоненты базы знаний ИС и соответствующего алгоритмического и программного обеспечения. Следующие этапы эффективной реализации измерительного процесса связаны с разработкой информационной технологии оперативного контроля состояния ДО и прогноза его поведения в различных условиях эксплуатации, в том числе и в экстремальных ситуациях. Конкретное наполнение знаниями прикладной области экспертизы определяется особенностями измерительного процесса и решаемых задач с учетом целей и наложенных ограничений.

4 Интеллектуальные технологии альтернативных решений на основе методов планирования эксперимента

Интеллектуальная поддержка планирования эксперимента в зависимости от особенностей исследуемой ситуации позволяет давать полезные советы по принятию решений на неформализованных этапах процесса экспериментального исследования, включая идентификацию ДО, выбор модели взаимодействия, выбор критерия оптимальности плана эксперимента, процедуры рандомизации и схемы эксперимента. В качестве модели представления знаний используется семантическая сеть. Ядром системы интеллектуальной поддержки является набор функциональных блоков, работающих со знаниями, хранящимися в семантической сети. В вершинах сети располагаются модели представления четких и нечетких знаний и необходимая информация для выполнения операций по планированию эксперимента. Компьютерная программа обеспечивает работу системы в виде последовательной активации вершин сети, содержащих списки ссылок на другие вершины.

База знаний интеллектуальной поддержки планирования эксперимента содержит различные типы знаний, среди которых выделяют знания описательного характера (факторы, источники неоднородностей и планы экспериментов) и логические правила, обеспечивающие принятие решений на основе анализа посылок, а также знания, содер-

жащие результаты компьютерного моделирования различных ситуаций при принятии решений о выборе плана эксперимента с учетом экономических, статистических и информационных требований. Механизм логического вывода поддерживается стеком активных вершин со списком номеров вершин, ожидающих очереди активизирования. База данных системы содержит данные, используемые в процессе логического вывода, и вспомогательные базы, включающие в себя каталог планов эксперимента, библиографическую информационную систему и каталог теорем. Программа для генерации планов эксперимента хранится на диске в отдельных файлах.

Моделирование текущих ситуаций обеспечивается путем решения задач взаимодействия судна с внешней средой, представленных в виде различных математических моделей, в том числе и блочно-модульной структурой. При этом исследуемая задача в заданной ситуации описываются в виде спецификации модели рассматриваемой области знаний. По спецификации аппарат предметного уровня формирует прикладную модель и перевод ее в контексте формального уровня модели предметной области, с помощью которого формируется модель ситуации и алгоритмическая ее реализация в виде набора прикладных программ.

5 Нечеткий логический базис при анализе конкурирующих вычислительных технологий методами планирования эксперимента

Рассмотренные в настоящем разделе методы планирования эксперимента предусматривают возможности использования достижений этой теории в различных задачах поддержки принятия решений. Однако в экспериментах, связанных с интерпретацией текущих ситуаций, особенно при сложном взаимодействии судна с внешней средой, возникают определенные трудности. *Именно поэтому новый подход к построению алгоритмов принятия решений должен предусматривать построение плана эксперимента в рамках принципа конкуренции, позволяющего выбирать предпочтительную вычислительную технологию*. Поскольку принцип конкуренции основан на развитии концепции мягких вычислений, при адаптации стандартных планов эксперимента приходится учитывать особенности нечетких и нейросетевых моделей, а также алгоритмы эволюционного моделирования.

Приведем примеры типичных логических правил, используемых при реализации принципа конкуренции.

1. *«ЕСЛИ число исследуемых факторов не превышает n=5 И все факторы представляют собой однородные экспериментальные единицы И данные эксперимента представлены нелинейной моделью, ТО рекомендуется использовать планы Коно или центральные композиционные планы»*.

2. *«ЕСЛИ опыты проводятся на неоднородных экспериментальных единицах И источники неоднородностей можно группировать в подмножества-блоки, ТО рекомендуются блочные рандомизированные планы»*.

Механизм логического вывода поддерживается стеком активных вершин со списком номеров вершин, ожидающих очереди активизирования.

Исходная математическая модель. Выбор исходной математической модели, описывающей динамику взаимодействия в нечеткой среде при планировании эксперимента, состоит в определении вида функции [1], [9]:

$$Y = f(x_j), (j=1,\dots,x_k). \quad (19)$$

Геометрический образ выходной характеристики *y* рассматривают как функцию или поверхность отклика, а функцию $f(x_j)$ принимают в виде отрезков степенных рядов (алгебраических полиномов) различных степеней. Функции $f_j(x)$ при j=1,2,...,k образуют матрицу известных коэффициентов F, которая полностью описывается выбором модели $f_j(x)$. По данным n измерений величины y_j для независимых переменных x_i (i=1,2,...,n) в принятой области планирования можно составить матрицу F и вектор Y результатов наблюдений:

$$F = \begin{bmatrix} f_0(x^1) & f_1(x^1) & \dots & f_k(x^1) \\ f_0(x^2) & f_1(x^2) & \dots & f_k(x^2) \\ & \dots & & \\ f_0(x^n) & f_1(x^n) & \dots & f_k(x^n) \end{bmatrix}. \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}. \tag{20}$$

Матрица F и вектор Y полностью описывают информацию для определения неизвестных коэффициентов математической модели, интерпретируемой в рамках теории нечетких множеств. Выбор числа и условий эксперимента, обеспечивающих получение наилучшего результата, составляет цель планирования эксперимента в нечеткой среде. В качестве критерия оптимальности в задачах планирования нечеткой системы используется D-*оптимальность*. Учет случайной и систематической погрешности производится на основе обобщенного критерия оптимальности плана X*, минимизирующего на множестве планов выражение

$$\int_{\Omega} E\left[\sum_{i=1}^{k} f_i(x)\theta_i - y(a,x)\right]^2 dx, \tag{21}$$

в котором в качестве модели $\sum_{i=1}^{k} f_i(x)\theta_i$ может быть выбран полином меньшей степени, чем в истинной модели.

Проблемы теории планирования эксперимента при интерпретации текущих ситуаций в задачах принятия решений связаны со сложностью процессов взаимодействия судна с внешней средой в условиях неопределенности и неполноты исходной информации. Процессы функционирования в условиях такого взаимодействия носят динамический характер, а измерения параметров определяются повышенными требованиями к их точности.

Оптимальная структура модели. Рассмотрим в качестве исходной математической модели зависимость [3], [5], [13]:

$$Y = X\beta + \varepsilon; \tag{22}$$

$$E(\varepsilon) = 0,\ \mathrm{cov}(\varepsilon) = \sigma^2 I, \tag{23}$$

где X – $(N\times k)$ – матрица наблюдений регрессоров x_j; β – вектор-столбец значений неизвестных коэффициентов размерности $(k+1)$; ε – вектор-столбец значений регрессионной ошибки размерности $(k+1)$; Y – вектор-столбец выходной переменной; E – оператор математического ожидания; I – единичная матрица; N – объем выборки; k – число регрессоров; σ^2 – дисперсия воспроизводимости.

Точность прогноза модели (22) определяется дисперсией ошибки предсказания. Для $(N+1)$ наблюдения по адекватной модели эта дисперсия равна

$$D\left[\Delta Y_{N+1}\right] = \sigma^2\left[X_{N+1,k}\left(X^T X\right)^{-1} X^T_{N+1,k+1}\right], \tag{24}$$

где $X_{N+1,k}$ – вектор-строка размерности $(1\times k)$.

Для сопоставления дисперсий моделей с «недобором» и «перебором» регрессионных переменных относительно истинной зависимости выражение (24) преобразуется к виду:

$$D\left[\Delta Y_{N+1}\right] = \sigma^2 [X_{N+1,k+1}\left(X_{k-1}^T X_{k-1}\right)^{-1} \times X_{N+1,k-1}^T + I] + \beta_k^2 \big/ \lambda_k^2 \,\hat{x}_{N+1,k-N+1,k}^2, \qquad (25)$$

где λ_k – параметр нецентральности t-распределения с N-k степенями свободы

$$\lambda_k = \left(\beta_k \big/ \sigma\right)\hat{\sigma}_k \left[(N-1)\left(1-\hat{R}_j^2\right)\right]^{1/2}; \qquad (26)$$

σ_k^2 – выборочная дисперсия переменной x_k; $\hat{R}_j$ – выборочный коэффициент множественной корреляции переменной x_k с остальными независимыми переменными $x_1, \dots, x_{k-1}$, вычисленный по данным матрицы наблюдений X.

Дисперсия ошибки прогноза с «недобором» и «перебором» будет равна

$$D\left[\Delta Y_{N+1}^*\right] = \sigma^2 [X_{N+1,k-1}\left(X_{k-1}^T X_{k-1}\right)^{-1} \times X_{N+1,k+1} + 1] + \beta_k^2 \left(\hat{x}_{N+1,k}^2 + x_{N+1,k}^2\right); \qquad (27)$$

$$D\left[\Delta Y_{N+1}^{**}\right] = \sigma^2 [X_{N+1,k-1}\left(X_{k-1}^T X_{k-1}\right)^{-1} \times X_{N+1,k-1}^T + 1] + \beta_k^2 / \lambda_k^2 \left(\hat{x}_{N+1,k+1} - x_{N+1,k+1}\right)^2. \qquad (28)$$

Здесь β_k – значение коэффициента при k-й неучтенной переменной x_k; $x_{N+1,k}$ – значение (N+1)-го наблюдения k-й переменной;

$\hat{x}_{N+1,k} = X_{N+1,k-1}\hat{A}_{k-1}$ – регрессия переменной $\hat{x}_k$ на остальные переменные $\hat{x}_k = \sum_q^{k-1} a_q x_q$; $\hat{A}_{k-1} = \left(X_{k-1}^T X_{k-1}\right)^{-1} X_{k-1}^T X_k$ – вектор-столбец коэффициентов размерности (k-1)×1.

На практике приходится избегать переусложнения модели за счет включения дополнительных регрессоров, так как

$$D\left[\Delta Y_{N+1}\right] \le D\left[\Delta Y_{N+1}^{**}\right], \qquad (29)$$

что следует из сопоставления (24) и (25). При этом равенство дисперсий достигается только при

$$\hat{x}_{N+!,k+1} = x_{N+1,k+1}. \qquad (30)$$

Выбор наилучшего уравнения поверхности отклика может быть произведен также с использованием базисных функций [3], [5] и модели вида

$$\eta = \sum_{k=1}^{K} \beta_k f_k(u), \qquad (31)$$

где β_k – k-й коэффициент регрессии; $f_k(u)$ – k-я базисная функция ($f_1(u) \equiv 1$); $u = (u_1, \dots, u_M)^T$ – вектор факторов пассивного эксперимента; M – число факторов; K – общее число слагаемых. При этом факторы кодированы, а модель наблюдения имеет вид

$$Y = \eta + \varepsilon. \qquad (32)$$

Системный анализ задачи выбора наилучшей полиномиальной регрессии позволяет сформулировать подход к определению степени m_{opt} алгебраического многочлена $P_m(c,x)$, достаточно надежно аппроксимирующего заданную экспериментальную зависимость $\{y_n,x_n\}$. Суть его состоит в дополнении классического метода (выбор в качестве m_{opt} величины m, доставляющей минимальное значение сумме квадратов отклонений) описанием способа измерения экспериментального массива $\{y_n\}$

$$y_n = g(G(x_n) + \varepsilon_n), \tag{33}$$

где $G(x_n)$ – некоторая функция; ε_n – случайные ошибки ($n = 1,N$).

Особенности вычислительной технологии. Вычислительные трудности при реализации алгоритмов обработки информации в трудноформализуемых средах, особенно данных пассивного эксперимента, связаны с процедурой обращения матрицы X^TX системы нормальных уравнений

$$(X^TX)^{-1}\beta = X^TY. \tag{34}$$

Как показывают результаты вычислений, обусловленность информационной матрицы X^TX существенно хуже, чем матрицы X. Если $\lambda_{\max}$ и $\lambda_{\min}$ – наибольшее и наименьшее сингулярные числа матрицы X, а $\lambda^2_{\max}$ и $\lambda^2_{\min}$ – матрицы X^TX, то при небольшом λ_N число λ^2_N может иметь порядок, сравнимый с порядком погрешностей. В результате матрица X^TX становится почти вырожденной, тогда как матрица X таким свойством не обладает. При решении плохообусловленных задач используют методы псевдообращения, обеспечивающие работу вычислительных процедур в условиях вырожденности. Псевдообратная матрица X^* существует для любой матрицы X, и для любого вектора $Y\hat{\beta} = X^*Y$ с минимальной нормой среди всех векторов, минимизирующих сумму квадратов отклонений $\left\| Y - X\hat{\beta} \right\|^2$.

Сравнительный анализ экспериментальных данных с помощью различных подходов обращения матриц позволяет выделить прямой метод Гревилла, позволяющий работать непосредственно с матрицей X. Эта матрица лучше обусловлена, чем матрица X^TX. В результате существенно повышается устойчивость решения. При использовании нелинейных по параметрам моделей возникает проблема оценки коэффициентов регрессии. При некомпактности априорного множества параметров такая оценка может вообще не существовать (не достигается минимум суммы квадратов отклонений). Алгоритмы минимизации суммы квадратов основаны на методе Ньютона – Гаусса и сводятся к линейной аппроксимации функции регрессии.

Планирование эксперимента в ИС исследовательского проектирования связано с выбором оптимальных условий, обеспечивающих организацию процедурной компоненты базы знаний и соответствующего алгоритмического и программного обеспечения. Конкретное наполнение знаниями прикладной области определяется структурой исследуемого процесса взаимодействия и номенклатурой решаемых задач с учетом поставленных целей и наложенных ограничений.

Практические рекомендации. Приведем некоторые рекомендации по использованию методов планирования эксперимента в ИС исследовательского проектирования. Численный пример оценки критериальных соотношений для различных способов формирования регрессионной структуры приведен в табл. 1. Здесь $D[DY_{N+1}]$ – дисперсия ошибки; w – частота; R_j – выборочный коэффициент множественной корреляции; S_0 – остаточная сумма квадратов; 1 – полная адекватная модель; 2 – модель с «недобором»; 3 – модель с «перебором».

Таблица 1 – Характеристики различных способов формирования регрессионной структуры

Вид модели Оценка	Рассматриваемые критерии			
	$D[DY_{N+1}]$	W	R_j	S_0
1 1	0,27	0,85	0,90	$0,46 \cdot 10^{-2}$
2	0,38	0,76	0,73	$0,58 \cdot 10^{-2}$
3	0,35	0,78	0,80	$0,55 \cdot 10^{-2}$

В зависимости от характера исследования эксперимент можно выполнить либо в виде последовательной серии опытов, либо случайным образом (рандомизированный эксперимент). Типичными планами эксперимента в задачах исследовательского проектирования являются матрицы плана Коно и центрального композиционного плана [6], [13], наиболее часто используемых в задачах интерпретации ситуаций в сложной динамической среде. Для построения оптимального плана эксперимента необходимо задаться видом модели, определить область варьирования переменных и выбрать число испытаний n. Если полученное значение n кратно двум или четырем, то для построения оптимального плана можно использовать *матрицы Адамара*.

Если помимо линейных эффектов исследуемая модель содержит взаимодействие факторов, то при построении плана эксперимента с целью сокращения числа измерений удобно использовать *дробные факторные планы*, а в отдельных ситуациях – *неполноблочные* сбалансированные планы (квадраты Юдена). Если исследуемая функция неизвестна, то можно ограничиться применением стандартного плана, при котором последовательно варьируется каждая из независимых переменных при условии неизменности всех остальных.

Заключение

Для развития науки исследовательского проектирования характерно противоречивое единство двух тенденций: дифференциации и интеграции. Дифференциация обусловлена различием в объектах и методах теоретических и экспериментальных исследований, спецификой используемых научных дисциплин и приложений. Тенденции к интеграции знаний проявляются в росте значения науки исследовательского проектирования как единой системы знаний, в распространении общих методов исследования и их математизации. Развитие интеллектуальных технологий в условиях неопределенности и неполноты исходной информации требует эффективного использования методов планирования эксперимента в ИС исследовательского проектирования на базе высокопроизводительных средств обработки информации в сложной динамической среде.

Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 – 2013 годы (ГК № П295 от 30.04.2010, ГК № П976 от 27 мая 2010, ГК № 02.740.11.0837 от 11.06.2010). Результаты работ использованы в рамках проекта реализации Правительства РФ № 218 «О мерах государственной поддержки развития кооперации российских высших учебных заведений и организаций, реализующих комплексные проекты по созданию высокотехнологичного производства».

Литература

1. Нечеткие множества в моделях управления и искусственного интеллекта / [Аверкин А.Н., Батыршин А.Н., Блишун А.Ф. и др.] ; под ред. Д.А. Поспелова. – М. : Наука, 1986.
2. Бортовые интеллектуальные системы. Ч. 1. Авиационные системы. Ч. 2. Корабельные системы. – М. : Ра-

диотехника, 2006. Ч. 3. Системы корабельной посадки летательных аппаратов. – М. : Радиотехника, 2008.
3. Горский В.Г. Планирование промышленных экспериментов / Горский В.Г., Адлер Ю.П., Талалай А.М. – М. : Металлургия, 1978.
4. Емельянов В.В. Теория и практика эволюционного моделирования / Емельянов В.В., Курейчик В.В., Курейчик В.Н. – М. : Физматлит, 2003.
5. Кастнер С. Планирование измерительного эксперимента в интеллектуальных системах реального времени / С. Кастнер, Ю.И. Нечаев, А.Б. Дегтярев // Сб. докл. Международной конференции по мягким вычислениям и измерениям SCM-2000. – Санкт-Петербург, 2000. – Т. 2. – С. 69-74.
6. Налимов В.В. Логические основания планирования эксперимента / В.В. Налимов, Т.И. Голикова. – М. : Металлургия , 1981.
7. Нейрокомьютеры в интеллектуальных технологиях XXI века. – М. : Радиотехника, 2012.
8. Нечаев Ю.И. Математическое моделирование в бортовых интеллектуальных системах реального времени / Ю.И. Нечаев // Труды 5-й всероссийской научно-технической конференции «Нейроинформатика – 2003». – М. :МИФИ, 2003. – Лекции по нейроинформатике. Часть 2. – С. 119-179.
9. Нечаев Ю.И. Теория катастроф: современный подход при принятии решений / Нечаев Ю.И. – Санкт-Петербург : Арт-Экспресс, 2011.
10. Постон Т. Теория катастроф / Т. Постон, И. Стюарт. – М. : Мир, 1980.
11. Системы искусственного интеллекта в интеллектуальных технологиях XXI века. – Санкт-Петербург : Арт-Экспресс, 2011.
12. Федоров В.В. Теория оптимального эксперимента / Федоров В.В. – М. : Наука, 1971.
13. Планирование эксперимента в исследовании технологических процессов / [Хартман К., Лецкий Э., Шеффер В. и др.]. – М. : Мир, 1977.
14. Thom R. Catastrophe theory: Its present state snd future perspectives / R. Thom // Dynamical systems. Warwick. 1974. – Berlin – Heidelberg – New York : Springer Verlag. 1-75. P. 366-372. Lecture Notes Math. V. 468.
15. Zadeh L. Fuzzy logic, neural networks and soft computing / L. Zadeh // Commutation on the ASM-1994. – Vol. 37, № 3. – P. 77-84.

Literatura

1. Averkin A.N. Nechetkie mnozhestva v modeljah upravlenija i iskusstvennogo intellekta. M.- Nauka. 1986.
2. Bortovye intellectual’nye sistemy. Ch. 1. Aviacionnye sistemy. Ch. 2.Korabel’nye sistemy. – M.: Radiotehnika. 2006. Ch. 3. Sistemy korabel’noj posadki letatel’nyh apparatov. M.: Radiotehnika. 2008.
3. .Gorskij V.G. Planirovanie promyshlennyh jeksperimentov. Metallurgija. 1978.
4. Emel’janov V.V. Teorija i praktika jevoljucionnogo modelirovanija. M.: Fizmatlit. 2003.
5. Kastner S., Sb. dokl. Mezhdunarodnoj konferencii po mjagkim vychislenijam i izmerenijam SCM-2000. Sankt-Peterburg. 2000. T.2. S.69-74.
6. Nalimov V.V. Logicheskie osnovanija planirovanija jeksperimenta. M.: Metallurgija. 1981.
7. Nejrokom’jutery v intellektual’nyh tehnologijah HHI veka. M.: Radiotehnika, 2012.
8. Nechaev Ju.I. Trudy 5-j vserossijskoj nauchno-tehnicheskoj konferencii “Nejroinformatika – 2003”. M.:MIFI.2003. Lekcii po nejroinformatike. Chast’ 2. S.119-179.
9. Nechaev Ju.I. Teorija katastrof: sovremennyj podhod pri prinjatii reshenij. Sankt-Peterburg: Art-Jekspress. 2011.
10. Poston T. Teorija katastrof. M.: Mir.1980.
11. Sistemy iskusstvennogo intellekta s intellektual'nyh tehnologijah HHI veka. Sankt-Peterburg: Art-Jekspress. 2011.
12. Fedorov V.V. Teorija optimal’nogo jeksperimenta. M.: Nauka. 1971.
13. Hartman K. Planirovanie jeksperimenta v issledovanii tehnologicheskih processov. M.: Mir. 1977.
14. Thom R. Catastrophe theory: Its present state and future perspectives. Dynamical systems. Warwick. 1974. Berlin – Heidelberg – New York Springer Verlag. 1-75. P. 366-372. Lecture Notes Math. V. 468.
15. Zadeh L. Sommutation on the ASM-1994. Vol.37. №3. P. 77-84.

Статья поступила в редакцию 05.06.2012.